DOCUMENT RESUME

ED 121 185 HE 007 623

AUTHOR Nunn, Richard; Lain, Lindy
TITLE Classification of Medical Education Institutions.
INSTITUTION Association of American Medical Colleges, Washington, D. C.
SPONS AGENCY Health Resources Administration (DHEW/PHS), Bethesda, Md. Bureau of Health Manpower.
PUB DATE Dec 75
NOTE 87p.; Prepared by Division of Operational Studies
AVAILABLE FROM Association of American Medical Colleges, Division of Operational Studies, One Dupont Circle, N.W., Washington, D.C. 20036

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage
DESCRIPTORS *Cluster Grouping; Factor Analysis; Grouping Procedures; *Higher Education; Literature Reviews; Medical Education; *Medical Schools; Research Methodology; *School Surveys; *Statistical Data
IDENTIFIERS *Institutional Profile System

ABSTRACT

Empirical techniques are developed that may be used in conjunction with data stored in the Institutional Profile System to enhance present capabilities of assessing group structure in medical schools. Relevant literature is reviewed, and the institutionally descriptive data available for analysis and their manipulation into researchable formats are described. In order to relate the present data and methods to previous studies, variables similar to those used by the RAND Corporation were chosen. The data are factor analyzed, and the factor scores then used in two empirical cluster analysis procedures, and the results compared to those generated by RAND. Lack of replication is noted. The 10 clusters of medical institutions found by RAND were not found in this study. A variety of factors may have contributed to this conclusion, including a substantial number of both methodological and data differences between the studies. It is concluded that, at least for the present, categorization of medical schools by procedures accounting for multiple measures simultaneously does not yield clear and unambiguous results. This finding indicates that the picture presented by data institutionally descriptive of the schools is a highly complex one, not easily structured into a reasonably small number of groups of institutions. (LBH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED121185

CLASSIFICATION OF MEDICAL EDUCATION INSTITUTIONS

PREPARED BY:

RICHARD NUNN
STAFF ASSOCIATE

AND

LINDY LAIN
RESEARCH ASSISTANT

ASSOCIATION OF AMERICAN MEDICAL COLLEGES
DIVISION OF OPERATIONAL STUDIES
ONE DUPONT CIRCLE, N.W.
WASHINGTON, D.C. 20036

HE 007 623

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

DECEMBER 1975

CLASSIFICATION OF MEDICAL EDUCATION INSTITUTIONS

The work upon which this publication is based was supported in part by the Bureau of Health Manpower, Department of Health, Education and Welfare pursuant to contract number 231-75-0007. However, any conclusions and/or recommendations expressed herein do not necessarily represent the views of the supporting agency.

Chapter V - DISCUSSION AND CONCLUSION
by Douglas J. McRae, Ph.D.
Senior Staff Associate
Division of Operational Studies

Graphics by Debra Pitzer
Division of Operational Studies

TABLE OF CONTENTS

		Page No.
I.	INTRODUCTION	1
	A. Review of the Literature	2
	B. Overview to the Current Study	5
II.	METHODS	6
	A. Hierarchical Clustering Schemes	9
	1. Illustration	11
	2. Ward's Objective Function Method	17
	B. Non-Hierarchical Clustering Schemes	19
	1. Illustration	20
	2. Forgy's K-Means Technique	22
III.	DATA	26
IV.	APPLICATION	33
	A. The RAND Study	33
	B. Replication: Factor Analysis	34
	C. Replication: Cluster Analysis	36
	1. Ward's Objective Function Method	39
	2. Forgy's K-Means Technique	44
	D. Comparison of Results	52
V.	DISCUSSION AND CONCLUSION	56
	A. Methodological Considerations	56
	B. Substantive Considerations	60
	C. Conclusion	62
	BIBLIOGRAPHY	65
	ABBREVIATIONS KEY	67
	APPENDIX A.	
	APPENDIX B.	

Chapter 1

INTRODUCTION

This report, addressing Article I, Section 1, Part b(1), of the "Series of Analytical Studies on Medical Education and Academic Health Centers" contract between the Bureau of Health Manpower of DHEW and the Association of American Medical Colleges, describes the development and application of several cluster analysis techniques to data descriptive of U.S. medical schools for the purpose of classifying the schools into several categories or groups. The tasks set forth in the contract are as follows:

(a) A general classification methodology shall be developed to identify ... parameters which are maximum in available data and which contain commonalities or discontinuities across institutions.

(b) The methodological approach shall focus upon developing analytic clustering methods useful for classifying institutions on the basis of any type of empirical data. Such a scheme shall provide a series of classifications corresponding to the particular subsets of data used. Data subsets so classified will include ... faculty mix, student mix, and output characteristics.

(c) The classification structures will then be cross-validated against any other quantifiable data representing component information, other published research in the field, and major positions by the medical education community, if available.

(d) Submit developed methodology in the form of a report.

[illegible] [illegible] [illegible] [

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Strictly interpreted, a key assumption of statistically based clustering procedures is that all objects must be placed into one and only one cluster. These procedures partition the entire set of objects (medical schools) into mutually exclusive and exhaustive subsets. This partitioning may take place whether the data sets are completely random or highly structured (i.e., whether or not there really are natural groupings). Outliers (unique objects similar to no other object) are either included in a cluster with other objects or constitute clusters by themselves. This conceptual constraint should be kept in mind when interpreting the results of any statistically based clustering procedure.

The two approaches to cluster analysis used in this study are in one case "hierarchical" and in the other "non-hierarchical". Each approach is described first, in general terms that include an illustration of typical results and then, in specific terms that detail the methodology chosen for this study.

A. Hierarchical Clustering Schemes

Hierarchical cluster analysis schemes generally construct groups of objects through a progression of stepwise merges. Initially, each object is considered a cluster in and of itself. A determination is then made as to which two clusters are most similar, whereupon these two clusters are merged. The process is then repeated until no further merge is possible.

This process starts with n objects or clusters, yields n-1 clusters after the first merge, n-2 clusters after the second merge, etc., until only one cluster (containing all n objects) remains. Hierarchical clustering schemes falling into this general framework have been labeled "agglomerative" hierarchical cluster analysis techniques*.

A feature of hierarchical, as opposed to non-hierarchical methods, is that once objects are grouped together they may not be separated later in the process. This feature offers both an advantage and a disadvantage. The early decisions greatly reduce the number of possible merges or changes that may take place later, thus allowing greater efficiency in the procedure. However, it precludes adjustment or reversal of unfortunate merges which have taken place earlier in the process.

Specific hierarchical clustering techniques differ from each other primarily in the criterion used to determine the basis for admittance of objects into clusters. These hierarchical procedures are described in this context in sections one and two below. In either case, an index depicting the status of the merge process, as plotted against the existing number of clusters, will often be helpful in determining an optimal solution found somewhere between the two extremes of n clusters

---

* There are a number of hierarchical techniques which work in a similar but reverse manner. They begin with a single cluster, containing all objects, and proceed to successively segment clusters into smaller and smaller groups. Such techniques are called "divisive" hierarchical cluster analysis procedures. They are not used in the study described in this paper.

(one object per cluster) and one cluster containing all n objects.

1. Illustration

To illustrate hierarchical clustering, consider the agglomerative procedure called the diameter method by Johnson (1967). This method fits into a general class of methods known as complete linkage. In the diameter method, the basic data analyzed is an n x n matrix of euclidian distances, where n is the number of objects to be clustered. At the first stage of clustering, the two objects with the smallest distance separating them are grouped together. At the next and all succeeding stages, an object is added to a cluster (note that a cluster may consist of only one object) only if the distance between it, the candidate object, and all objects within a cluster is less than its distance to all objects not in that cluster.

For example, four objects have the following distance matrix:

	A	B	C	D
A	0	1.0	2.0	1.5
B		0	0.5	2.5
C			0	3.0
D				0

This matrix indicates that the distance between object A and object B is 1.0 units, that the distance between object B and object D is 2.5 units, etc.

At the first stage in Johnson's diameter method, objects B and C are grouped together because they are separated by the smallest distance in the matrix (i.e., they are most similar). At the second stage, the distance between object A and object D (1.5 units) is smaller than the distance between A and C (2.0 units); therefore, neither A nor D may be added to the B-C cluster and are grouped together to form the second cluster. At the final stage, the two clusters (B-C and A-D) are grouped together to form one cluster containing all four objects.

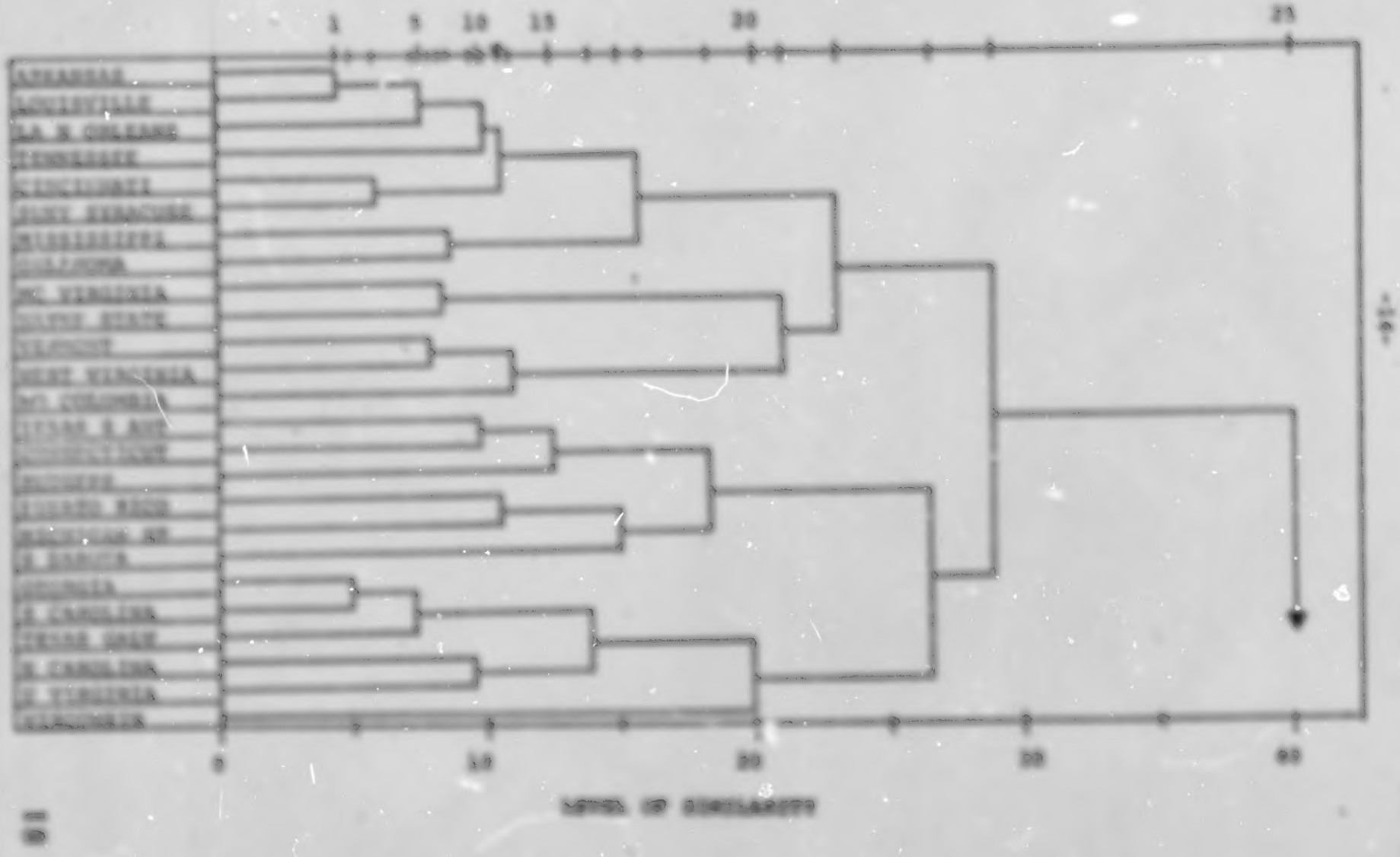
The merge criterion suggested by Johnson for this method is quite stringent and as a result produces clusters that are highly homogeneous. Although this characteristic may be beneficial under some circumstances, frequently, complete linkage methods are excessively constraining and fragmentary in their formulation of clusters. As Bailey summarizes: "complete linkage methods...dilate space. This means that the existing clusters move away from unclustered individuals as the clusters grow so that such individuals are more likely to form nuclei of new clusters than to add to pre-existing ones."

Despite this criticism, it is useful to examine the results produced by Johnson's diameter method because the method is simple and easy to interpret. To illustrate both the application of Johnson's diameter method and the typical graphic summary of results that normally accompanies hierarchical clustering, a dendrogram (tree diagram) is presented in Figure 2.1. Using data on each of 99 medical institutions for each of 6 factor scores (see Chapter IV for description of the factor scores), a 99 x 99 matrix of distances has been generated and submitted to analysis using the Johnson diameter method. Results shown in the figure are for 24 of the 99 institutions. The school names are listed on the left side of the page, and the critical distance for each merge is shown across the bottom of the figure. The sequence in which merges take place is recorded on the top of the figure.

A dendrogram may be interpreted by observing the development of linkages shown by series of interconnecting lines. Before a merge between any two schools takes place (at any point before merge sequence "1" is encountered), each school has a single line projecting horizontally to the right. At this stage there exists as many clusters as there are medical schools (25). The first merge, depicted by a link or vertical connecting line, occurs between Arkansas and Louisville. The distance separating these two schools can be determined by locating the corresponding position on the scale, "Level of

FIGURE 2.1

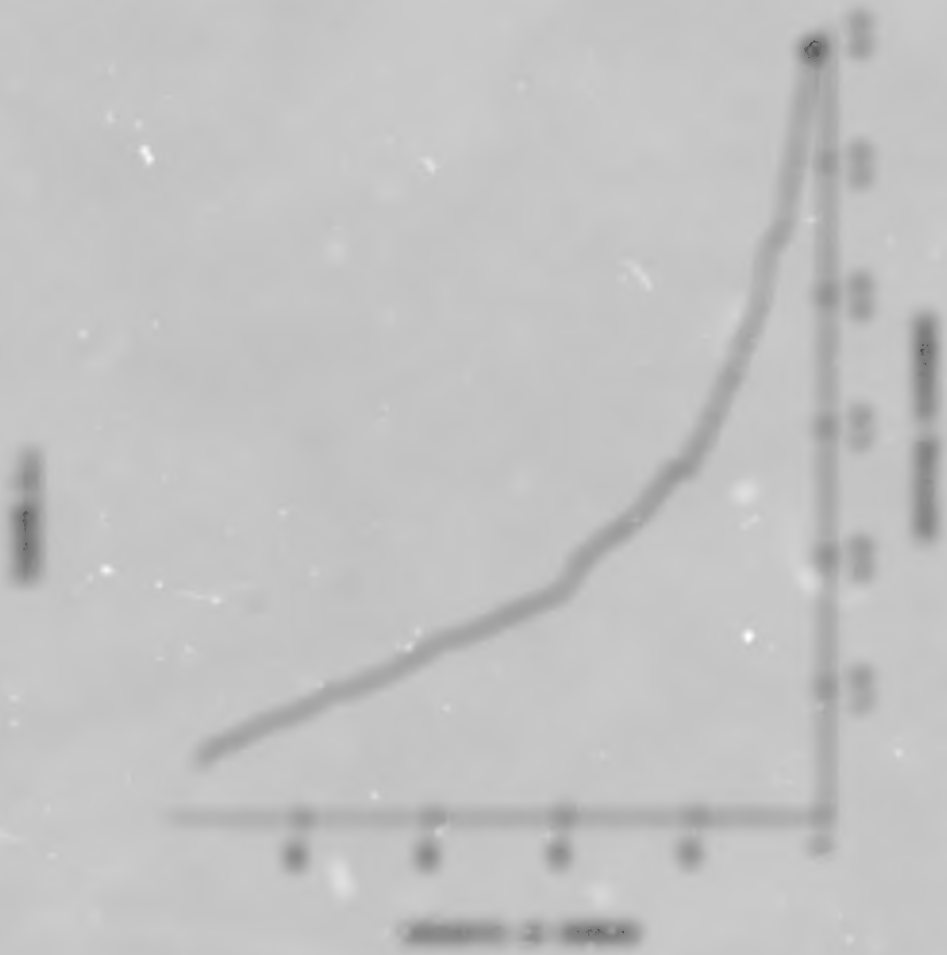
Dendrogram: Illustration of Hierarchical Cluster Formulation Based on Johnson's Diameter Method



[illegible]. In this case, Aylesmere and Lewisville are [illegible] as a unit and, as such, are no longer [illegible] by individual horizontal lines but with a single line reflecting their joint status as members of the same cluster. The same steps may be applied throughout the dendrogram through n-1 merges, until all national schools fall into one cluster. For purposes of illustration, a twenty-fifth merge has been included in the merge sequence table. At this point in the clustering process, all 25 national schools linked in the present dendrogram merge to a single group with another such group to form one cluster.

Since hierarchical techniques potentially present all cluster solutions, some guidance is needed in selecting an "optimal" solution, i.e. to answer the question: "How many clusters are there?" This determination will often be more apparent when plotting the number of clusters existing at any given stage against the critical distance (or distance criterion) is used in the merging process). Such a plot, for the full 99 school analysis using the Ward's "distance" method, is given in Figure 3.3.

This plot allows one to weigh the benefits of combining clusters against the sacrifices in group cohesion, expressed as critical distances, needed to facilitate a merge. The reduction in the number of clusters from 99 to 85, for example, incurs only a slight relaxation in the critical distance. At the other extreme, merging into progressively fewer clusters exacts



[illegible] the [illegible] [illegible] [illegible]

[illegible]

[illegible] $H_{2}O$ [illegible]

[illegible]

[illegible]

[illegible]

a final pass through the data is made in order to reassign objects that have become closer to other centroids. The final pass requires no further updating of the centroids.

As originally proposed by MacQueen, K-means is a two pass procedure. The first pass, just described, finds centroids; a second pass makes final assignments of objects to clusters. It is important to note that objects assigned to clusters on the first pass may be assigned to different clusters on the second pass. As indicated above, this is one property that distinguishes non-hierarchical techniques from hierarchical techniques.

It should also be noted that by assigning objects to clusters based upon smallest distances, the K-means technique is very similar to the Ward method which attempts to minimize the within-group sum of squares. MacQueen (1967) presents theoretical and empirical evidence of this similarity. Thus, the K-means procedure and Ward's objective function procedure share similar objectives but differ primarily in the arbitrary procedures designed to achieve these objectives.

Forgy (1967) suggests modification to MacQueen's basic K-means method in two substantial ways. First, he suggests that the process continue iterating as long as an objective function, such as the within-groups sum of squares, continues to decrease. Secondly, he suggests that centroids not be

recalculated until the end of each iteration. These two changes result in a non-hierarchical process similar to the Ward method in its formal attempts to minimize an objective function. The Forgy modifications, however, overcome two of the major problems of the Ward procedure, the permanence of cluster membership inherent in the hierarchical approach and the difficulties associated with migrating means.

In all non-hierarchical procedures, the specification of initial cluster centroids (seed points) is of great importance. This specification may be done randomly as in MacQueen's method in which the first g objects are taken, or it may be based upon some substantive grounds. In particular, if the investigator has some notion of where concentrations of objects occur in the structure, he may wish to ensure that clusters be given every chance to grow in that area. Thus, one modification to the Forgy procedure that may be used, if sufficient knowledge of the data structure is available, is that of specifying values to serve as initial seedpoints.

One further observation, although both the Ward technique and Forgy technique attempt to minimize the within-group sum of squares, there is no guarantee that either technique will reach the absolute minimum. The only way to ensure the attainment of an absolute minimum sum of squares is through complete enumeration of the data set, but even aided by today's

advanced computer technology, a complete enumeration is unrealistic for all but very small data sets.

The two clustering procedures chosen for the present study are used in successive stages. First, a hierarchical cluster analysis, Ward's objective function, is used. Second, results from Ward's method provide seed points for analysis of the same data by Forgy's non-hierarchical procedure. Computer programs for both of these procedures have been obtained and adopted for use on the AAMC's Institutional Profile System. The use of these procedures is illustrated in Chapter IV. The procedures are now available for application to subsets of data in the Institutional Profile database.

Chapter III

DATA

Beyond variations evident in the objectives defined by clustering methodologies, the properties assumed in any configuration of clusters is essentially a reflection of the data employed. This chapter focuses on the development of a data set suitable for use with cluster analysis. The two areas to be considered are, first, the availability and selection of variables from the Institutional Profile System and, second, the preparation of the data for analysis.

The Institutional Profile System (IPS) is a computerized information retrieval system with a large data storage capacity and software to perform various statistical and data summary functions. Currently, the data base includes data on U.S. medical schools for 16 years and contains over five thousand variables from 49 source questionnaires. The sources of interest for the present study are primarily the Liaison Committee on Medical Education Questionnaire: Parts I and II, 1973-74.

A. The Data Set: Availability and Selection

The principle objective in developing a data set was to assemble a comprehensive set of variables with a sufficiently broad, yet detailed, perspective of medical education to facilitate exploratory analyses. In addition to this study,

two other exploratory studies under this contract also required utilization of such a data set. (Sherman, 1975; Sedlacek, 1975). Basically, then, the variables in the data set are intended to have a general rather than contextual descriptive value.

The contract specification related to this study, provides for the development of cluster analytic methods that may be used as an additional function within the IPS statistical package. Because, however, the cluster algorithms made available to the organization are not compatible with the program language used in IPS, all analyses conducted for this paper were done external to IPS. The first step, then, involved extraction of relevent data from IPS in order that these analyses might be performed. The studies conducted by Sherman and Sedlacek also required an external application of the data for use in the statistical programs available in SPSS • (Nie, 1975).

Variable selection involved the identification of the most current and meaningful data available. Selection began with the most current IPS data for 117 medical institutions. Although the bulk of the data was available for academic year

---

• Statistical Package for the Social Sciences

1973-74, the most current financial data was for 1972-73.

In attempting to formulate a full spectrum of salient institutional descriptors, suggestions were elicited from RAND staff representing a number of specific areas in academic medicine. Other potentially useful descriptors were noted in the studies by Richards (1966), Rodgers, Otis, and RAND and were organized into logical domains given in Table 3.1. To facilitate comparison, variables used by each study are identified by an X in the columns to the right. The emphasis of each study on particular variable domains, as shown in this table, varies considerably. The most obvious omissions occur in the faculty domain, with the exception of RAND, and in the curriculum domain for all but Otis' study.

The more current and extensive data available in IPS allow an expansion of the data set for the present study to 350 variables. The set contains approximately 235 variables taken directly from source documents and an additional 130 variables derived from the original 235, (mostly ratios and percents). The entire set is listed in Appendix A. A summary of the expanded domains is given in Table 3.2.

B. Data Preparation

The development of a data set, particularly one of this size, requires a number of preliminary tasks. These tasks include organizational considerations, such

[illegible]

I. [illegible]

A. [illegible]

1. [illegible]
2. [illegible]
3. [illegible]
4. [illegible]
5. [illegible]
6. [illegible]
7. [illegible]

A. [illegible]

1. [illegible]
2. [illegible]
3. [illegible]
4. [illegible]
5. [illegible]
6. [illegible]
7. [illegible]
8. [illegible]
9. [illegible]

A. [illegible]

1. [illegible]
2. [illegible]
3. [illegible]
4. [illegible]
5. [illegible]
6. [illegible]
7. [illegible]
8. [illegible]
9. [illegible]
10. [illegible]
11. [illegible]
12. [illegible]
13. [illegible]
14. [illegible]
15. [illegible]
16. [illegible]
17. [illegible]

II. [illegible]

A. [illegible]

1. [illegible]
2. [illegible]
3. [illegible]

A. [illegible]

A. [illegible]

1. [illegible]
2. [illegible]
3. [illegible]

A. [illegible]

1. [illegible]
2. [illegible]

[illegible]

[illegible] 2.2

[illegible] (Rs. [illegible])

1.	[illegible]	[illegible]
	a. [illegible]	[illegible]
	b. [illegible]	[illegible]
	c. [illegible]	[illegible]
2.	[illegible]	[illegible]
	a. [illegible]	[illegible]
	b. [illegible]	[illegible]
	c. [illegible]	[illegible]
	d. [illegible]	[illegible]
	e. [illegible]	[illegible]
3.	[illegible]	[illegible]
	a. [illegible]	[illegible]
	b. [illegible]	[illegible]
4.	[illegible]	[illegible]
	a. [illegible]	[illegible]
	b. [illegible]	[illegible]
5.	[illegible]	[illegible]
	a. [illegible]	[illegible]
	b. [illegible]	[illegible]
	c. [illegible]	[illegible]
	d. [illegible]	[illegible]
	e. [illegible]	[illegible]
	f. [illegible]	[illegible]

[illegible]

RAND, that of common factor extraction and equimax rotation, Sherman finds essentially the same six factors based on 117 institutions as opposed to RAND's 94. They are, as presented in Table 4.3: (1) graduate medical education programs, (2) Federal research involvement, (3) undergraduate medical education programs, (4) reliance on non-full-time faculty, (5) public versus private control and (6) non-M.D. education programs.

A factor, which may be viewed as simply a synthetic variable, is a condensation of a group of variables into a single expression. Each school's position along the descriptive dimension represented by a "factor" (such as the second, "Federal research involvement") is given a "factor score" computed from the input variables using a formula derived by the factor analysis. There are six such formulae, one for each factor. Each school, then, has six factor scores that replace values for the 23 vairables. The computed similarity of two schools is a composite measure of the similarity of their six factor scores. This composite measure includes a set of numerical weights to reflect the subjective importance of each of the six dimensions in determining the school's similarity. The present study uses the same numerical weights assigned by RAND.

TABLE 4.3

Factor Pattern Matrix from Analysis of
RAND Study Variables (Using New AAMC Data)
By Method of Common Factors and Equimax Rotation

VARIABLES	RAND FACTOR LABELS	VARIABLE LABELS	FACTOR (VARIABLE GROUPS) 1	2	3	4	5	6
V6310	Graduate Medical Education Programs	TOT RESDNTS INSTR BY MD FAC 73-74	.86	.22	.28	-.03	.07	-.03
V6330		TOT INTERNS INSTR BY MD FAC 72-73	.80	.19	.21	-.02	.05	-.00
V6080		ENROL RATIO-INTERNS & RESDNTS TO MD STU	.64	-.10	-.29	.09	.11	.22
V6010		TOT STUDENTS...ALL...INSTRUCTED AT MC	.61	.38*	.60	-.17	-.12	.21
V3350	Federal Research Involvement	SPONS PROG EXPD PER FT FAC	.10	.86	.19	-.00	.11	.09
V3345		MC EXPD-REG OP COSTS PER MD STUDENT	.24	.61	-.19	.50	.09	.17
V2830		MC EXPD-PCT SPONS PROG EXPD OF TOT	.04	.67	.25	-.06	.34*	.12
V2940		NIH AWARDS RESRCH GRANTS $1000 73-74	.42*	.58	.06	.35*	.30	.36*
V7200		MEAN MCAT SCORE SCI-1ST YR MD STUDENTS	.38*	.41	.00	.09	.23	.30
V2820		PCT SPONS FAC SALARY FROM FED $ 72-73	-.14	-.32	-.12	.02	-.20	.26
V6020	U.G. Med. Educ. Programs	ENROLL-TOT MD STUDENTS 73-74	.30	.22	.85	-.08	.02	-.11
V1045		AGE OF INSTITUTION	.08	.08	.74	-.07	.29	.04
V1140		SMSA POP PER MD STUDENT	.17	.06	-.33	.11	-.05	-.27
V5025	Reliance on Non-Full-Time Faculty	RATIO FT FAC TO TOTAL STUDENTS	-.11	-.02	-.29	.76	.05	.04
V2750		TOT MC EXPD PER TOTAL STUDENTS	-.19	.52	-.12	.81	.06	-.07
V5040		RATIO VOL FAC TO FT FAC	.05	.02	-.45*	-.41	-.16	-.37
V2740		SPECIAL PROJ $ PER MD STUDENT 72-73	-.20	.11	-.11	-.34	-.00	.01
V1030	Control: Public Vs. Private	CONTROL TYPE (1=PRIVATE, 0=PUBLIC)	.10	.05	.01	.07	.78	-.10
V6230		PCT MD STUDENT FROM HOME STATE	.11	-.20	-.24	.16	-.58	-.09
V2110		MC REV-TOT UNRESTR ENDOW & GIFTS	.14	.17	.06	.29	.56	.11
V6050	Non-M.D. Educ. Programs	ENROLL RATIO-MD STUDENT EQUIV TO MD STU	.08	.12	.16	.24	-.29	.64
V6140		ENROLL RATIO-MAS & DOC BAS SCI TO MD STU	.17	.16	.00	.07	.07	.52
V5030		RATIO PT FAC TO FT FAC	.03	.01	.03	-.30	-.20	-.43

C. Replication: Cluster Analysis

1. Analysis by Ward's Objective Function Method

Based on the six factors identified by Sherman, two independent cluster analyses are conducted. The first cluster application, which is to be discussed in the immediate section, is the Ward objective function method. The application of the Forgy method follows in Section 2.

Although Sherman's factor analysis replication employs 117 medical institutions, 99 have been retained for cluster analysis. Data for the other 18 schools is missing for four (18%) or more of the 23 original vairables. Of the 99 medical schools used in this study, 83 are in common with the 94 used by RAND.

The particular variation of Ward's technique used in this analysis (Wishart, 1969) requires a symetric matrix of euclidean distances. In other words, a distance is computed for each pair of schools to reflect a composite of their differences on the six factor scores. All such combinations are stored in a 99 x 99 matrix and used as input for analysis by the Ward method.

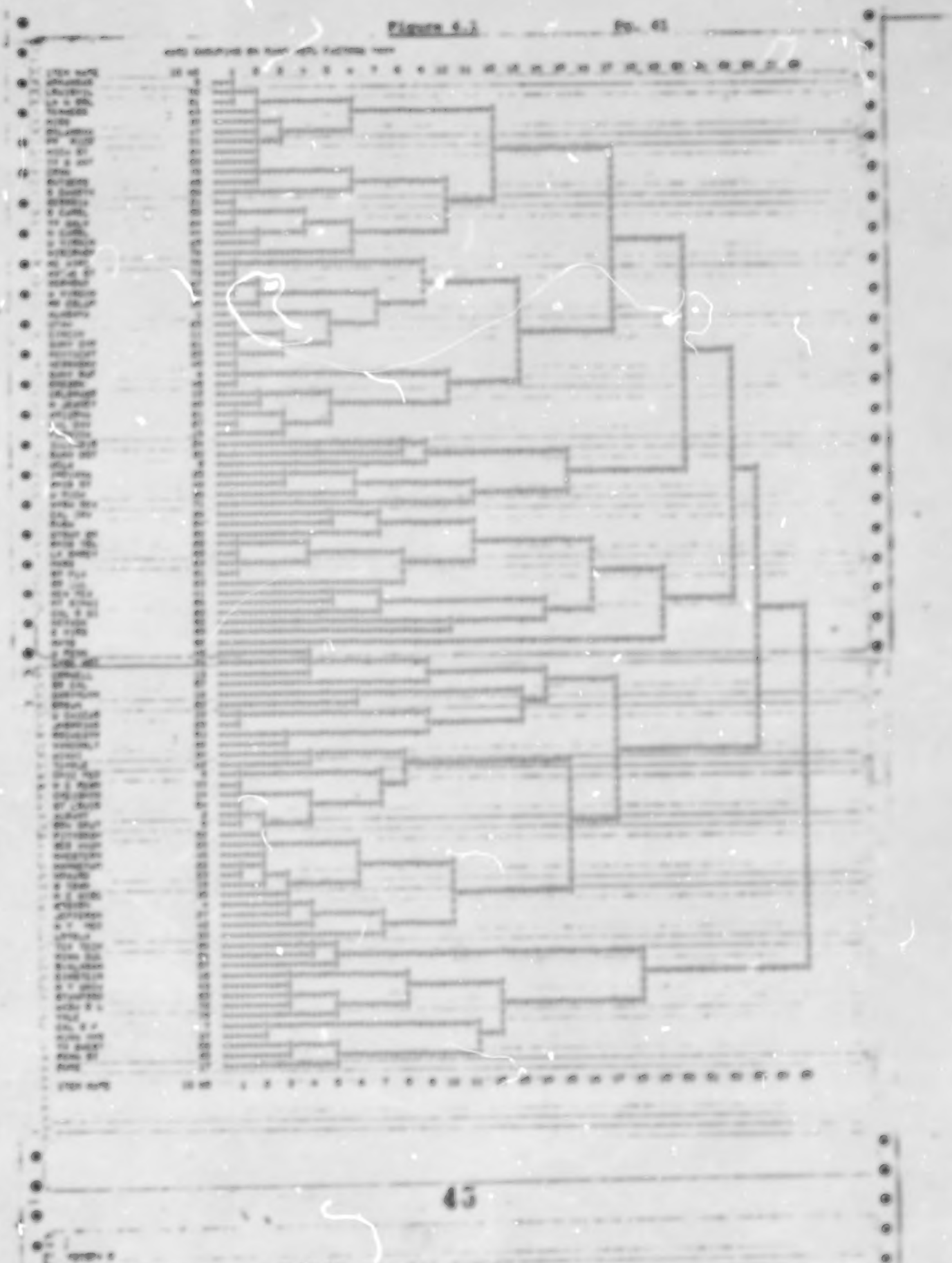
The result of the Ward analysis is provided in the form of a dendrogram in Figure 4.1. The dendrogram is condensed in order to reflect stages of the merge process rather than the full 98 individual (n-1) merges. These steps are equally incremented to preserve the actual error sum of squares

needed to accomplish each merge. The 99 institutions analyzed are in the left colurn of the figure and the sequence number for 25 merge stages is shown across the top.

Figure 4.1 indicates that Arkansas, Louisville, and Louisiana-New Orleans merge at stage 1, that Tennessee joins this cluster at stage 2, and that this cluster does not admit any new members until stage 6 when a cluster consisting of Mississippi, Oklahoma, and Puerto Rico joins it.

On a broad perspective, the dendrogram reveals an elongated pattern of cluster growth. Merges between schools appear to occur fairly uniformly throughout the procedure; however, during the earlier stages they branch laterally for some distance before expanding vertically to admit larger numbers of institutions. This trend indicates that group structure is, on the whole, relatively introspective and confined to small memberships.

To illustrate this point, note that clusters forming up to stage 5 are numerous but generally contain only two to four members. In fact, at level 5 there are 44 clusters, or an average of only 2.25 schools per cluster. The largest cluster contains six schools, while a total of 17 schools have not as yet merged. If, for example, a single cluster is to be formed from the first 18 institutions listed in the dendrogram, the criterion level is found to be "12." The internal structure of this cluster in turn is made up of three subgroups of roughly

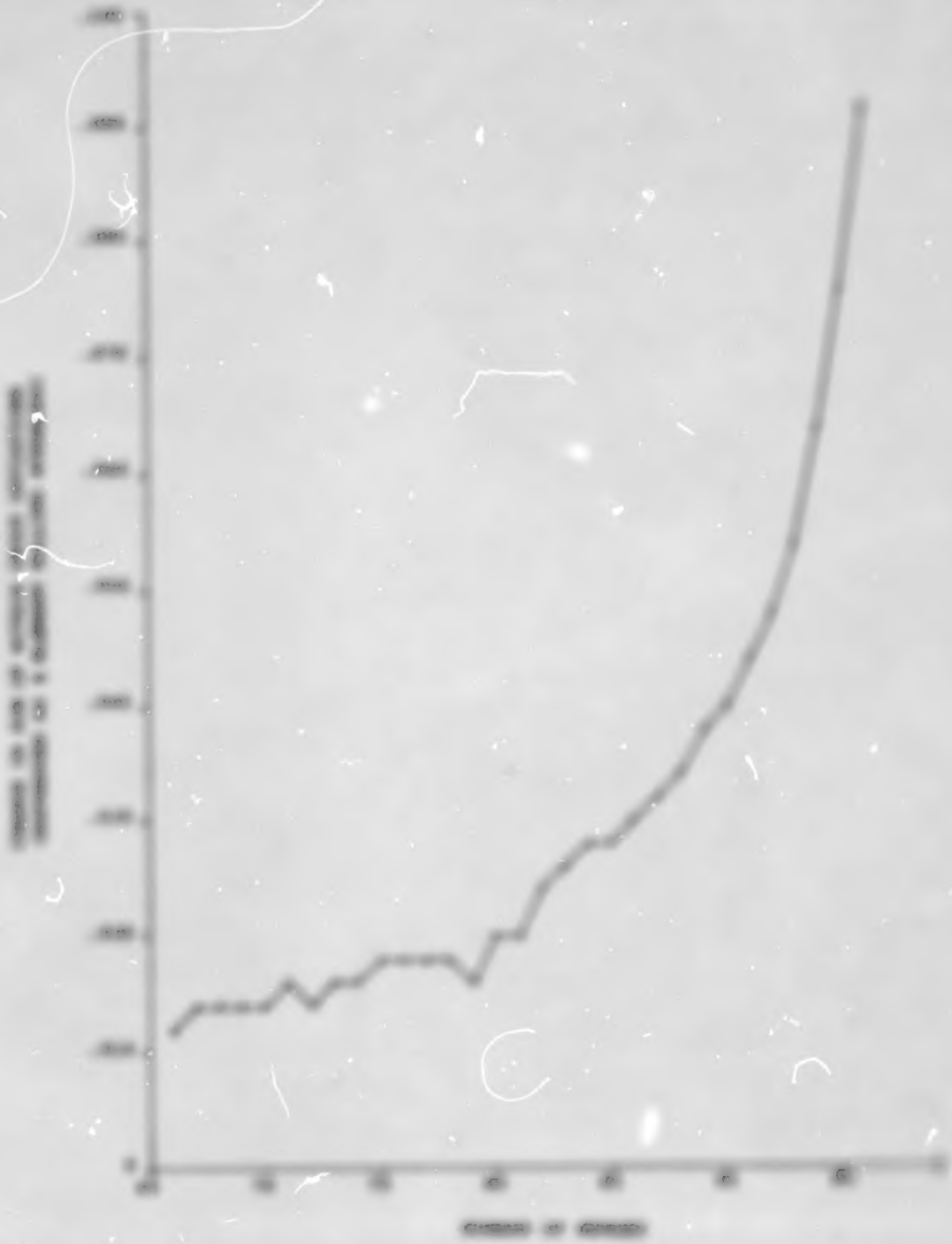


equal size. The criterion level has to be doubled, from level 6 to level 12, in order to tolerate the merging of these three subclusters.

In an earlier example, Figure 3.3, a plot of the number of clusters existing at any given stage against the critical distance has been suggested as a guide in selecting an optimal solution. It is perhaps more meaningful in the present context to plot the change in the sum of within-group deviations incurred at each merge. Additionally, to reduce the effects of local extremes, change in the sum of deviations is expressed as five element rolling averages. The resulting plot is given in Figure 6.2.

A plot of this kind often reveals disproportionate jumps transpiring between group deviations and the progression of merges. Figure 6.2, for instance, reveals perhaps two transitional points in the curve. The first point is at the eighty-first merge where 10 clusters have been formed; the second, at the ninetieth merge with nine clusters. These two transitional points suggest two optimal solutions. Schools grouped in the 10 cluster and 9 cluster solutions are given in Appendix D.

Ultimate determination of the "optimal" number of clusters is, of course, primarily subjective. Such determination must take into account the nature of the data analyzed, the methods used, and the particular goals of the analysis. The purpose of using hierarchical clustering

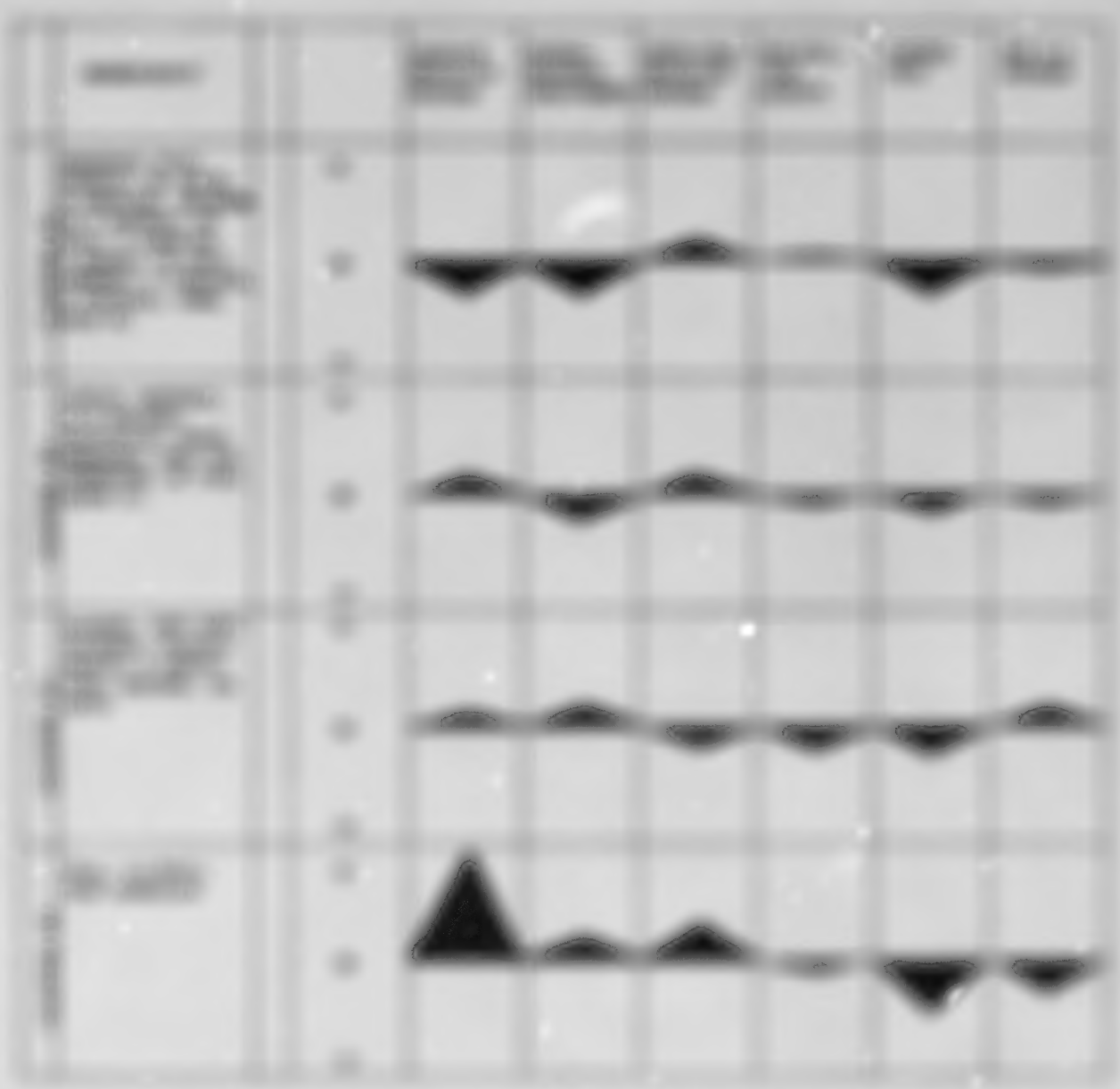


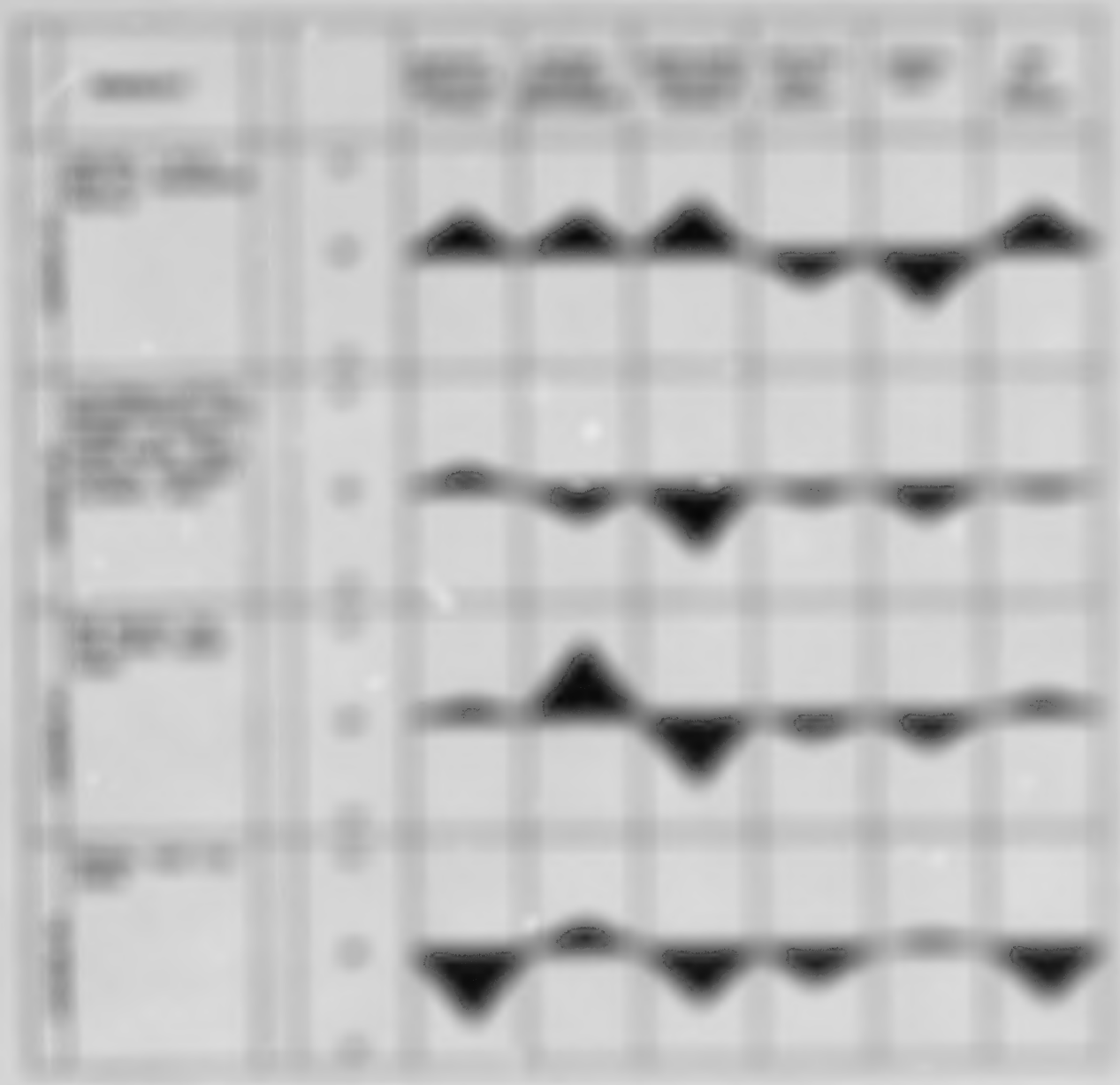
[illegible] [illegible] [illegible] [

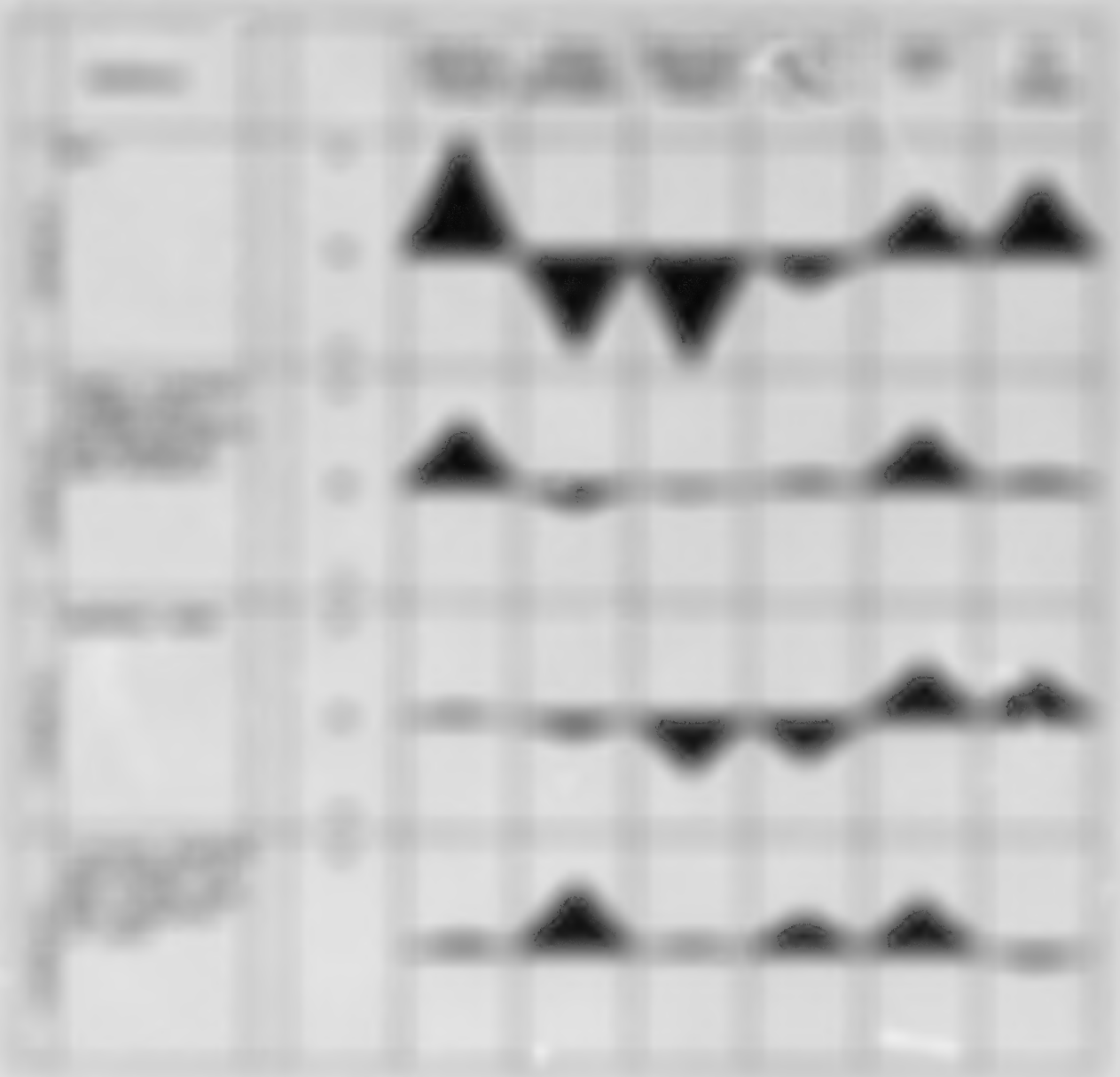
[illegible]

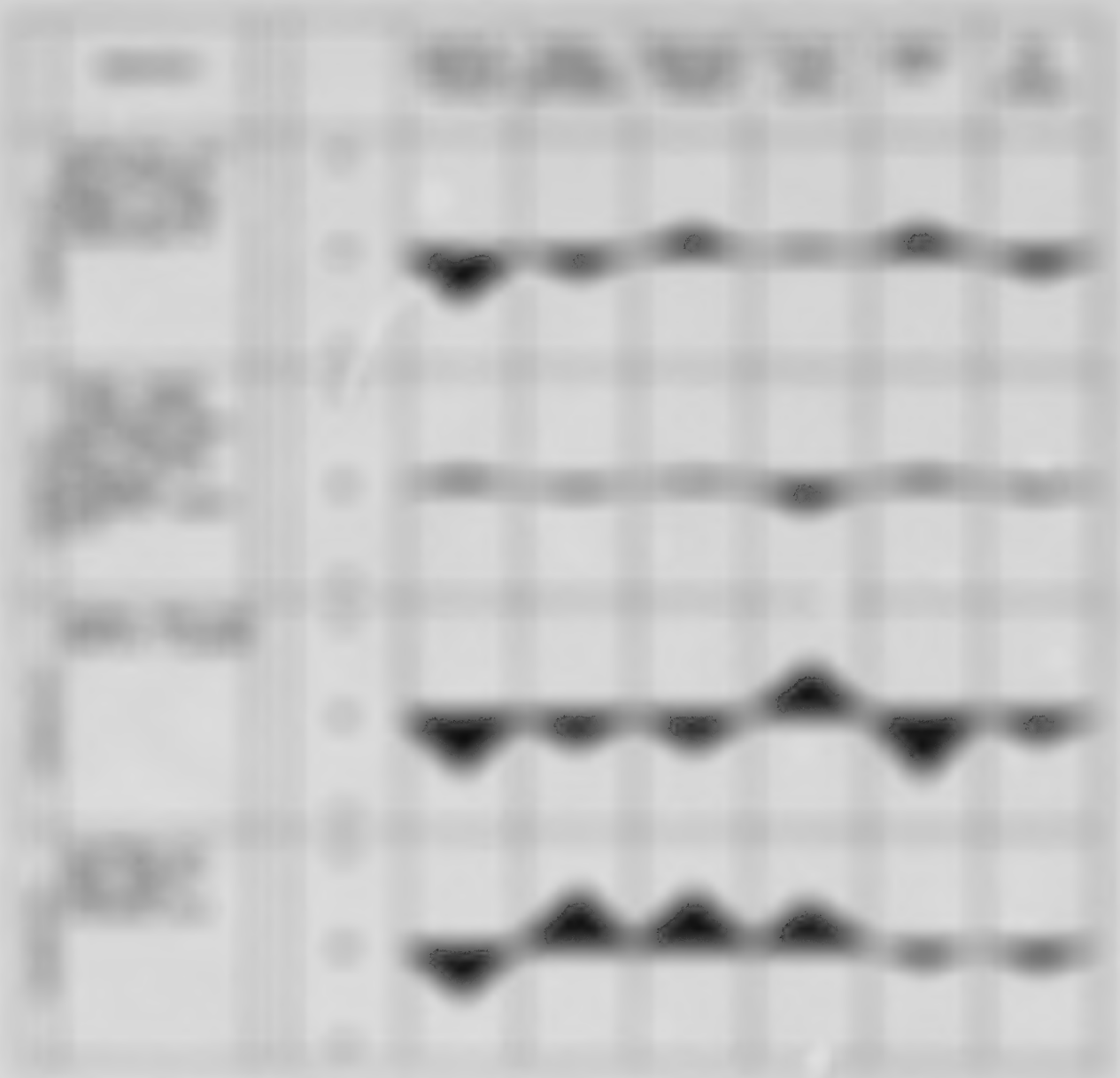
[illegible]	[illegible]	[illegible]					
		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

[illegible]









ment in both federal research and undergraduate education, yet are somewhat below the mean in their reliance on non-full-time faculty and in non-M.D. programs offered. The four medical schools in cluster 5 have much the same attributes with two exceptions. First, the graduate education program is smaller and, second, the non-M.D. programs offered by these schools is comparatively high.

The differences between clusters 10 and 11 lie primarily in the intensity of scores on the six factors. First, the five schools in cluster 10 and two schools in 11 are all private. Both clusters have a lower than average commitment to federal research. Cluster 10 has a rather sizeable graduate education program with an average undergraduate education program and reliance on non-full-time faculty. The profile of cluster 11 on the three factors is, comparitively speaking, somewhat lower.

The profiles of the remaining five clusters are not directly comparable because their profiles approach varying extremes. Two of these clusters may well be considered outlying groups. Cluster 9, for instance, contains only Mayo. It can be readily seen from its profile that Mayo is quite unlike any other medical school in that it offers very large graduate and non-M.D. programs with an absence of involvement in federal research and undergraduate programs.

The two medical schools grouped in cluster 8, Nevada and Eastern Virginia, are new schools and as one would expect fall into lower extremes. In terms of research involvement, however, cluster 8 is slightly above the mean.

Cluster 15 is also made up of new medical schools. As a group, they differ from cluster 8 in their reliance on non-full-time faculty.

Cluster 12 consists of a group of all private schools distinguishable by heavy research involvement and reliance on non-full-time faculty. In all other regards, this cluster exhibits average characteristics.

A group of six public schools form cluster 6. These schools, besides having relatively extensive graduate programs, are otherwise on the lower extreme of the spectrum of variables.

Finally, the profile of cluster 16, which includes a mixture of public and private schools, indicates a lesser emphasis on graduate and non-M.D. educational programs, though the emphasis on undergraduate programs is substantial. Furthermore, the schools do have a significant involvement in federal research and reliance on non-full-time faculty.

D. Comparison of Results

A relatively simple way of comparing group membership between two sets of clustering results is to equate the number of matches and mismatches in tabular form, such as

shown in Table 4.5. In the left hand column may be found the cluster identification numbers assigned by RAND to its resulting clusters. The cluster identification numbers shown across the top of the table are those evolving from the Forgy application to the replicated factors. For ease of interpretation, columns have been rearranged so as to reflect results of the two studies in corresponding order. In this way, clusters that are most highly associated in terms of membership fall roughly into cells along the diagonal.

Of those medical schools assigned by RAND to cluster 1, for example, consistency with the replicated cluster findings is minimal. The fifteen members of RAND's cluster 1 occur in small groups ranging across half the spectrum of the clusters derived from the replicated version. With few exceptions, much the same lack of conformity exists throughout. Notably, the exceptions are RAND's cluster 5, and to a lesser extent, 8 and 9. Members belonging to clusters 8 and 9 account for the bulk of private schools and differ mainly in regard to relative wealth and research orientation. The characteristics of the membership of cluster 5 are less obvious. Generally, the medical schools belonging to this cluster are public institutions with relatively low federal research involvement and graduate program emphasis.

TABLE 4.5

AAMC Cluster Numbers
As Assigned to Forgy Results

RAND CLUSTER NUMBERS	5	6	2	3	4	10	1	7	13	12	14	16	8	9	11	15	Outlier
1	2	2	3	3			2					2					1
2				1			1	1									
3	1		2		2	1					3						2
4	1				1	3				2		1					2
5			2				9					1					1
6		1		1				1									
7			3	2		1	2		3		1						
8										4	2	2					2
9									5	1	5						2
10				2				1									1

As stated earlier, the number of clusters formed in the current analysis has been increased beyond RAND's ten to 16 in order to provide for the effects of imposing a fixed number of clusters on differing populations. Specifically, the concern has been to allocate a sufficient number of clusters to enable medical schools not in the RAND Study to develop into groups external to RAND's ten. As such, the several new medical schools forming the replicated clusters 8, 9, 11 and 15 are completely absent from the RAND analysis. The schools belonging to these four clusters then are seperated out, thereby making the respective solutions more readily comparable. As a further note, no apparent combination of the 12 remaining replicated clusters enhances the over all conformity with the RAND clusters without, in turn, obscuring other equally important group distinctions.

Chapter V

Discussion and Conclusion

The present paper reports on two aspects of an attempt to utilize rather sophisticated statistical procedures to shed light on the "structure" of medical education in the U.S. The first aspect is primarily methodological, and it involves the descriptions in Chapter II and Chapter III regarding multivariate cluster analysis procedures and manipulation of available data into formats amenable to analysis by such procedures. The second aspect is primarily substantive, and it involves the description in Chapter IV regarding the generation of a classification of medical schools based on the methods and data developed, and comparison of these results to early efforts.

The discussion of the results reported, then, logically falls into two areas: methodological and substantive. The methodological aspect involves two components: choice of analysis procedure and constructing of a researchable data set. The present chapter is organized along these lines, with a final section devoted to conclusions that may be drawn from the report.

A. Methodological Considerations

1. Cluster Analysis Procedures

The cluster analysis procedures chosen for the present work are relatively new techniques in any data analyst's toolbox. As such, not a great deal is known about differences

to be expected when one set of methods is used in contrast to some other set of methods. The lack of "maturity" of these methods contributes to the generally unexplained phenomena that use of such methods and interpretation of results is still more art than science.

Statistically based clustering procedures first received wide scale attention from applied research methods scholars in the mid- and late 1960's. A number of procedures predate this time frame, but these generally did not receive either wide scale attention nor use. With the wide availability of high speed electronic computers for scientific applications in the mid-60's, however, iterative approximation procedures became feasible, and as a result a wide number of statistically based clustering analysis procedures were suggested. Unfortunately but not surprisingly, at this stage of the game little is known about the effects of using particular methods in contrast to other methods.

The state of the art of statistically based clustering procedures may be contrasted to the state of the art of factor analysis. Much of the conceptual work in factor analysis procedures was done in the 1930 to 1950 time frame. In the 1960's, differences that could be expected in application of one procedure vs. another were delineated, and several widely accepted "standard" procedures

[illegible] [illegible] [illegible] [

[illegible]

BIBLIOGRAPHY

Anderberg, M.R. Cluster Analysis for Applications. New York: Academic Press, 1973.

Bailey, K.D. "Cluster Analysis." Sociological Methodology 1975. David Hess (Ed.) San Francisco: Jossey-Bass, 1974.

Forgy, E.W. Cluster Analysis of Multivariate Data: Efficiency Versus Interpretability of Classifications. Biometric Society Meetings, Riverside, CA (Abstract in Biometrics 21:3,768), 1965.

Gower, J.C. "A Comparison of Some Methods of Cluster Analysis." Biometrics, 23:628-637, Dec., 1967.

Greenhouse, S.W. "On the Meaning of Discrimination, Classification, Mixture, and Clustering in Statistics." The Role and Methodology of Classification in Psychiatry and Psychopathology. M.M. Katz, J.O. Cole, and W.E. Barton, (Ed.). Chevy Chase, MD. U.S. Dept. of HEW, 1967.

Hartigan, J.A. Clustering Algorithms. New York: Wiley and Sons, 1975.

Jancey, R.C. "Multidimensional Group Analysis." Austrailian Journal of Botany, 14:127-130, 1966.

Jardine, N. and Sibson, R. "The Construction of Hierarchic and Non-hierarchic Classifications." Computer Journal, 177-184, Aug., 1968.

Johnson, S.C. "Hierarchical Clustering Schemes." Psychometrika, 32:241-254, Sept., 1967.

Keeler, E., Koehler, J.E., Lee, C., and Williams, Jr., A.P. Finding Representative Academic Health Centers. A Working Note prepared for NIH/HEW. Santa Monica, CA: Rand Corp., 1972.

Lance, G.N. and Williams, W.T. "Computer Programs for Hierarchical Polythetic Classification." Computer Journal, 9:60-64, March, 1966.

_____ "A General Theory of Classificatory Sorting Strategies: Hierarchical Systems." Computer Journal, 9:373-380, Feb., 1967.

McRae, D.J. "MIKCA: A FORTRAN IV Iterative K-Means Cluster Analysis Program" Behavioral Science, 16:423-424, 1971.

McQuitty, L.L. "Hierarchical Linkage Analysis for the Isolation of Types." Educational and Psychological Measurement, 20:55-67, Spring, 1960.

Mulaik, S.A. The Foundations of Factor Analysis. New York: McGraw-Hill, 1972.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., and Bent, D.H. Statistical Package for the Social Sciences. Second Edition. New York: McGraw-Hill, 1975.

Otis, G.D., Graham, J.R., and Thacher, L. "Typological Analysis of U.S. Medical Schools." Journal of Medical Education, 50:328-338, 1975.

Rodgers, S.A. and Elton, C.F. "An Analysis of the Environment of Medical Schools." Research in Higher Education, 2:239-249, Sept., 1974.

Rummel, R.J. "Understanding Factor Analysis." Journal of Conflict Resolution. 2:444-480, March, 1967.

Sedlacek, W.E. Variables Related to Increase in Medical School Class Size. Washington, D.C.: Association of Americal Medical Colleges, 1975.

Sells, S.B. "Toward a Taxonomy of Organizations." New Perspectives in Organizational Research. W.H. Cooper, H.J. Leavitt, and M.W. Shelly (Eds.) New York: John Wiley, 1964.

Shaw, K.E. "An Application of Cluster Analysis to the Study of Organization in a College of Education." Higher Education, 2:343-356, Aug., 1973.

Sherman, C.S. Study of Medical Education: Interrelationships Between Faculty, Curriculum, Student and Institution Variables. Washington, D.C.: Association of American Medical Colleges, 1975.

Tryon, R. and Bailey, D.E. Cluster Analysis. New York: McGraw-Hill, 1970.

Ward, Jr., J.H. "Hierarchical Grouping to Optimise an Objective Function." Journal of American Statistical Association, 58:236-244, 1963.

Wishart, D. "An Algorithm for Hierarchical Classification." Biometrics, 22:165-170, 1969.

ABBREVIATIONS

ACADM	ACADEMIC
ADMISS	ADMISSIONS
ADMN & GEN	ADMINISTRATIVE & GENERAL
ADV DEGREE	ADVANCED DEGREE
ADVIS PROG	ADVISORY PROGRAM
AFFIL	AFFILIATED
AM	AMERICAN
AMBUL	AMBULATORY
AMT	AMOUNT
ANESTH	ANESTHESIOLOGY
APPL	APPLICANT, APPLICATION
ASSOC PROF	ASSOCIATE PROFESSOR
ASSOC PROF MD	ASSOCIATE PROFESSOR OF MEDICINE
ASSTD	ASSISTED
AV	AVERAGE
BACH	BACHELORS DEGREE
BAS SCI	BASIC SCIENCE
BEHAV OBSS PUBLSHD	BEHAVIORAL OBJECTIVES PUBLISHED
BLDG	BUILDING
CL SCI	CLINICAL SCIENCE
CONSTR	CONSTRUCTION
CURR	CURRICULUM
DEPT	DEPARTMENT
DEV	DEVELOPMENT
DOC	DOCTORATE
DOC CAND	DOCTORAL CANDIDATE
DOC CONFRD	DOCTORALS CONFERRED
ED	EDUCATION
ENDOW	ENDOWMENTS
ENTERING	ENTERING STUDENTS
EQUIP	EQUIPMENT
EXPD	EXPENDITURES
FAC	FACULTY
FED	FEDERAL
FMS	FOREIGN MEDICAL STUDENTS
FT FAC	FULL-TIME FACULTY
GPA	GRADE POINT AVERAGE
GRAD	GRADUATION
GRTS	GRANTS

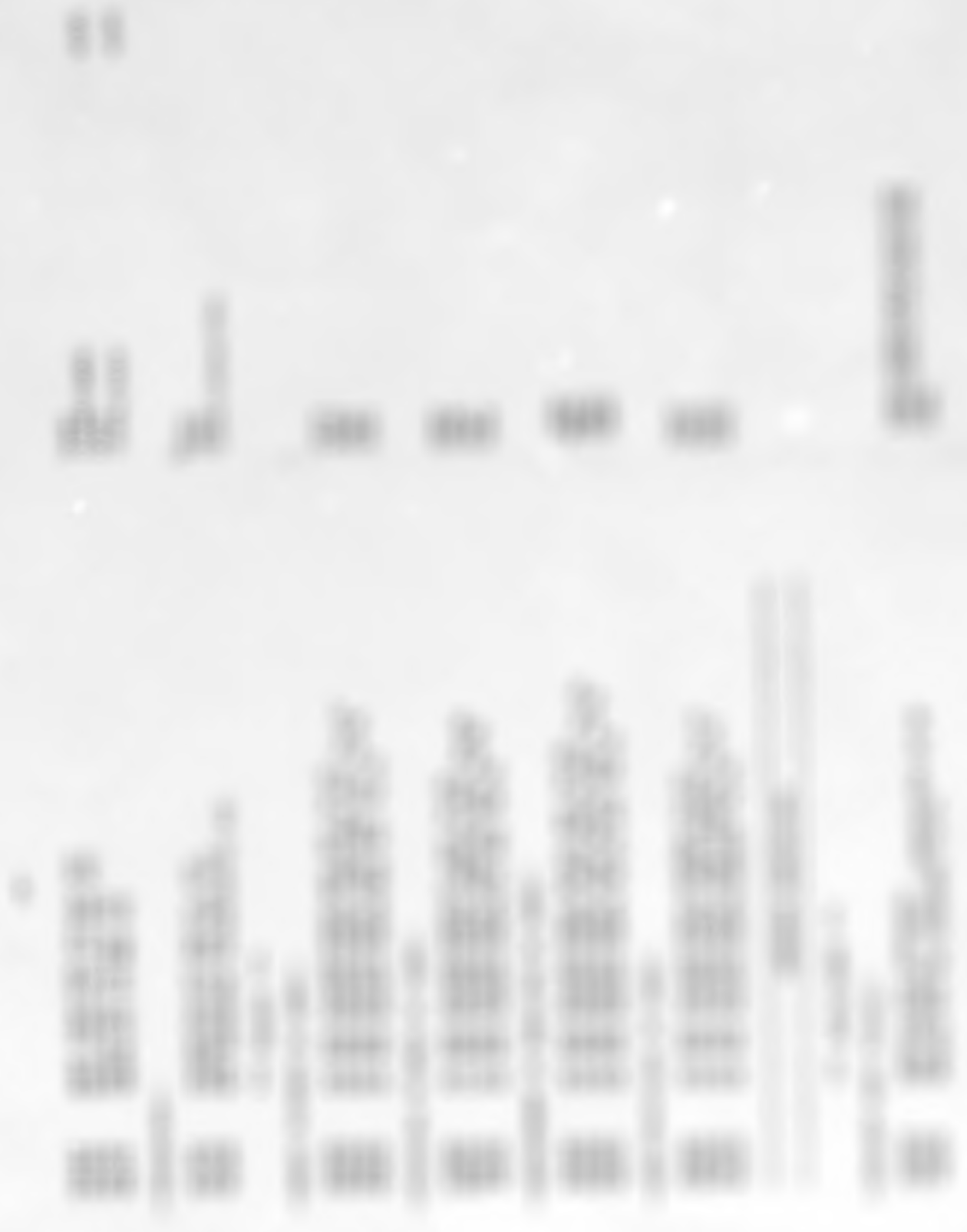
HLTH	HEALTH
HMO	HEALTH MAINTENANCE ORGANIZATIONS
HOSPS	HOSPITALS
HS SR	HIGH SCHOOL SENIOR
INDUS	INDUSTRY
INNOVATN	INNOVATION
INSTR	INSTRUCTOR, INSTRUCTED
INSTR & DEPT RESRCH	INSTRUCTION & DEPARTMENTAL RESEARCH
LOC	LOCAL
MANGMT	MANAGEMENT
MAS	MASTERS DEGREE
MC	MEDICAL COLLEGE
MCAT SCORE GEN	MCAT SCORE GENERAL KNOWLEDGE
MCAT SCORE SCI	MCAT SCORE SCIENCE
MCAT SCORE VER	MCAT SCORE VERBAL
MCAT SCORE QUAN	MCAT SCORE QUANTITATIVE
MD	MEDICAL
NATL BDS	NATIONAL BOARDS
NEED & RECVD AID	NEEDED & RECEIVED AID
NON-GOVT	NON-GOVERNMENT
PCT	PERCENT
PHYS ASST	PHYSICIAN'S ASSISTANT
POP	POPULATION
PRIM CARE	PRIMARY CARE
PRIV	PRIVATE
PROF	FULL PROFESSOR
PROF MD	PROFESSOR OF MEDICINE
PROG	PROGRAM
PROJ	PROJECT
PROJTD	PROJECTED
PT FAC	PART-TIME FACULTY
RECVD	RECEIVED
REG OP COSTS	REGULAR OPERATING COSTS
REQ AID	REQUESTED AID
REQ & RECVD AID	REQUESTED & RECEIVED AID
RESDNTS	RESIDENTS
RESRCH	RESEARCH
REV	REVENUES
REV CAREER	REVIEW CAREER

SCH	SCHOOL
SELECTD	SELECTED
SERV	SERVICE
SMSA	STANDARD METROPOLITAN STATISTICAL AREA
SPONS	SPONSORED
ST	STATE
STUDENT EQUIV	STUDENT EQUIVALENT
TCH-TRN	TEACHING-TRAINING
TOT	TOTAL
TRANS STUDENTS	TRANSFERRED STUDENTS
TUIT & EXPEN	TUITION & EXPENSES
UNIV	UNIVERSITY
UNRESTR	UNRESTRICTED
VOL	VOLUMES
VOL FAC	VOLUNTARY FACULTY
WITHDRL	WITHDRAWALS
YR	YEAR

APPENDIX A

VARIABLES LIST FOR CLASSIFICATION
OF MEDICAL INSTITUTIONS STUDY
AND INTERASSOCIATIONSHIPS STUDY

INSTITUTION	AHA SOURCE VARIABLES	AUTOMATICAL TABULATION
*** GENERAL CHARACTERISTICS***		
V1000 NO-IDENTIFICATION CODE	ID#	
V1010 STATE NO LOCATED	ST#	
V1020 REGION NO LOCATED	CEN	
V1030 CONTROL TYPE (1=NON-PUBLIC 2=GOVERNMENT)	CNTL	
V1040 YEAR FOUNDED	YRFD	
V1045 AGE OF INSTITUTION	YR INTw-YRFD	
V1050 1 IF 6 YR HOS	7005	
V1060 ACCREDITATION	2021	
V1070 LC TYPE 6 HOSPITAL	2847	
V1071 OBST APPL HOSPITAL	2847	
V1072 OBGY OR ANY APPL HOSPITAL	2847	
V1080 PSY BEDS APPL HOSPITAL	American Hospital Association Curriculum Elements	
V1085 RATIO APPL HOSP BEDS TO HO (1=YES)	American Hospital Association Curriculum Elements	
V1090 NUMBER OF BEDS APPRTD 72-74	Department of Institutional Development	
*** GEOGRAPHIC ***		
V1100 MC LOCATION-AREA FOR 71	CBSA	
V1110 MC LOCATION-CHARACTER LOCATION FOR 71	CBST	
V11— MC LOCATION-CHARACTER LOCATION FOR-ADDRESS 71	CBSA	
V1130 MC LOCATION-AREA FOR-(1=IN NON-AREA)	CBSA	
V1140 AREA FOR YES NO (1=YES)	CBSA/CBSI	
*** LIBRARY ***		
V1200 MC LIBRARIES-FOR YRS	2223	
V1210 MC LIBRARIES-ACCREDITATIONS	2224	
V1220 MC LIBRARIES-FOR SPECIAL TITLES BOOTS	2225	
FINANCES (NUMBERS) FOR 72-73		
*** REVENUE ***		
—REVENUE BY SOURCE—		
V2000 MC NON-FOR ALL SOURCES	CCOS	CCOS
V2010 MC NON-FOR PCG SOURCES	CLAS	CTAX
V2017 PCT OF MC REV FROM PAT SOURCES + PAYMENTS COST REVENUE	CCOS + CLAS/ CCOS	ACOS
—REVENUE BY SOURCE (GOVERNMENT)—		
V2100 MC NON-FOR GENERAL PURPOSE TITLES OTHR PLANS	CCOS	CCOS
V2110 MC NON-FOR GENERAL PURPS & CONTR	CNCS	ACOS
V2115 PCT OF TOT MC REV (1=> 10%)FOR GENOM & CONTR	CCOS/CCOS/CCOS	



V7535	REQ+RECVD AID-PCT 3RD YR MD STUDENTS	(1987/1977)	x100
V7540	REQ AID-TOT FINAL YR MD STUDENTS	1978	
V7545	REQ+RECVD AID-PCT FINAL YR MD STUDENTS	(1988/1978)	x100

--RECEIVING--

V7600	RECVD AID-TOT MD STUDENTS	1989	
V7610	TOT AID TO MD STUDENTS	1999	
V7615	AV AMT AID TO MD STUDENTS	(1999/1391)	x100
V7620	RECVD AID-TOT 1ST YR MD STUDENTS	1985	
V7630	TOT AID TO 1ST YR MD STUDENTS	1995	
V7635	AV AMT AID TO 1ST YR MD STUDENTS	1995/1985	
V7640	RECVD AID-TOT 2ND YR MD STUDENTS	1986	
V7650	TOT AID TO 2ND YR MD STUDENTS	1996	
V7655	AV AMT AID TO 2ND YR MD STUDENTS	1996/1986	
V7660	RECVD AID-TOT 3RD YR MD STUDENTS	1987	
V7670	TOT AID TO 3RD YR MD STUDENTS	1997	
V7675	AV AMT AID TO 3RD YR MD STUDENTS	1997/1987	
V7680	RECVD AID-TOT FINAL YR MD STUDENTS	1988	
V7690	TOT AID TO FINAL YR MD STUDENTS	199.	
V7695	AV AMT AID TO FINAL YR MD STUDENTS	1998/1988	

--NEEDING--

V7700	NEED AID-TOT MD STUDENTS	1984	
V7705	NEED+RECVD AID-PCT OF TOT MD STUDENTS	(1989/1984)	x100
V7710	NEED AID-TOT 1ST YR MD STUDENTS	1980	
V7715	NEED+RECVD AID-PCT 1ST YR MD STUDENTS	(1985/1980)	x100
V7720	NEED AID-TOT 2ND YR MD STUDENTS	1981	
V7725	NEED+RECVD AID-PCT 2ND YR MD STUDENTS	(1986/1981)	x100
V7730	NEED AID-TOT 3RD YR MD STUDENTS	1982	
V7735	NEED+RECVD AID-PCT 3RD YR MD STUDENTS	(1987/1982)	x100
V7740	NEED AID-TOT FINAL YR MD STUDENTS	1983	
V7745	NEED+RECVD AID-PCT FINAL YR MD STUDENTS	(1988/1983)	x100

--AID DISPERSED TO STUDENTS--

V7800	AID-AMT PER MD STUDENT	1999/1391	
V7810	RECVD AID-LOANS-TOT MD STUDENTS	2036	
V7815	RECVD AID-LOANS-PCT MD STUDENTS	(2036/1391)	x100
V7820	RECVD AID-SCHLSHIP-TOT MD STUDENTS	2037	
V7825	RECVD AID-SCHLSHIP-PCT MD STUDENTS	(2037/1391)	x100

*** EXPENSES ***

--TUITION, EXPENSES, & FEES--

V7900	TUIT+EXPEN PER IN ST MD STUDENT	1965	
V7910	TUIT+EXPEN PER OUT ST MD STUDENT	1966	
V7920	FEES+EXPEN EXCLUD TUIT PER MD STUDENT	1969	
V7930	AV EXPEN PER IN ST MD STUDENT UNMARRIED	2039	
V7940	AV EXPEN PER OUT ST MD STUDENT UNMARRIED	2043	
V7950	TUIT+EXPEN RATIO-IN ST TO OUT ST	1965/1966	

*** STUDENT SELECTION ***

--YEAR--

V8000	YR SELECTD-HS SR 73	1280
V8010	YR SELECTD-UNDERGRAD FR 74-75	1281
V8020	YR SELECTD-UNDERGRAD SOPH 74-75	1282
V8030	YR SELECTD-UNDERGRAD JR 74-75	1283
V8040	YR SELECTD-UNDERGRAD SR 74-75	1284

--APPLICANTS--

V8100	APPL-TOT	Division of Student Studies
V8110	APPL-TOT MALE	Division of Student Studies
V8115	APPL-PCT MALE TO TOT	Division of Student Studies
V8120	APPL-TOT FEMALE	Division of Student Studies
V8130	RATIO-MALE APPL TO ENTERING	Division of Student Studies
V8140	RATIO-FEMALE APPL TO ENTERING	Division of Student Studies
V8150	RATIO-APPL TO ENTERING	Division of Student Studies

--STANDING--

V8200	MC ACCEPT TRANS STUDENTS	1286
V8210	MC ACCEPT ADV STANDING STUDENTS	1311

*** CAREER REVIEW ***

V8300	REVIEW CAREER CHOICE AT GRADUATION	0438
V8310	REVIEW CAREER CHOICE 5 YRS AFTER GRAD 73	0439
V8330	ADVIS PROG-STUDENT RETENTION 74-75	1318
V8340	CAREER INTENT AFFECTS ADMISS DECISION	0441

APPENDIX B

19 CLUSTER SOLUTION		9 CLUSTER SOLUTION
CLUSTER #1	ARKANSAS LOUISVIL LA N ORL TENNESS MISS OKLAHOMA PR RICO MICH ST	CLUSTER #1
CLUSTER #2	TX S ANT CONN RUTGERS S DAKOTA GEORGIA S CAROL TX GALV N CAROL U VIRGIN WISCONSIN	
CLUSTER #3	MC VIRG WAYNE ST VERMONT W VIRGIN MO COLUM ALABAMA UTAH CINCIN SUNY SYR KENTUCKY NEBRASKA	CLUSTER #2
CLUSTER #4	SUNY BUF OREGON COLORADO N JERSEY ARIZONA CAL DAV FLORIDA	
CLUSTER #5	ILLINOIS SUNY DST UCLA	CLUSTER #3
CLUSTER #6	INDIANA OHIO ST U MICH WASH SEA	
CLUSTER #7	CAL IRV RUSH STONY BRK OHIO TOL LA SHREV MASS SO FLA SO ILL	CLUSTER #4
CLUSTER #8	NEW MEX MT SINAI CAL S DI	
CLUSTER #9	NEVADA E VIRGINIA	
CLUSTER #10	MAYO	CLUSTER #5

APPENDIX B

19 CLUSTER SOLUTION		9 CLUSTER SOLUTION
CLUSTER #11	U PENN CASE WST CORNELL SO CAL	
CLUSTER #12	DARTMOUTH BROWN	CLUSTER #6
CLUSTER #13	U CHICAGO J HOPKINS ROCHESTER VANDERBILT	
CLUSTER #14	MIAMI TEMPLE CHICAGO MED M C PENN CREIGHTON ST LOUIS ALBANY BOWMAN GRAY PITTSBURGH	CLUSTER #7
CLUSTER #15	GEO WASH NWESTERN HAHNEMAN HOWARD G TOWN M C WISCONSIN BOSTON JEFFERSON N Y MED LOYOLA	
CLUSTER #16	TEX TECH MINN DUL S ALABAMA	CLUSTER #8
CLUSTER #17	EINSTEIN F Y UNIV STANFORD WASH S L YALE	CLUSTER #9
CLUSTER #18	CAL S F MINN MPS	
CLUSTER #19	TX SWEST PENN ST DUKE	

Studies in Medical Education

Anderson, P. Descriptive Study of Salaried Medical School Faculty. December, 1975.

Johnson, D.G. and Dube, W.F. Descriptive Study of Medical School Applicants, 1974-75. December, 1975.

Lambdin, J.A. Survey of How Medical Students Finance Their Education, 1974-75. December, 1975.

Nunn, R. and Lain, L. Classification of Medical Education Institutions. December, 1975.

Rosenthal, J. Medical School Programs, Resources and Financing.

Sedlacek, W.E. Variables Related to Increases in Medical School Class Size. December, 1975.

Sherman, C. Study of Medical Education: Interrelationships Between Component Variables. December, 1975.

Additional copies of these publications may be obtained from:

Association of American Medical Colleges
Attention: Membership and Subscriptions
1 Dupont Circle N.W.
Washington, D.C. 20036